Fall 2018

Pupillometry as a test of infant word recognition

Amy Vinyard

Pupillometry as a Test of Infant Word Recognition

Amy Vinyard

A Thesis Submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Science

Department of Communication Sciences and Disorders

December 2018

FACULTY COMMITTEE:

Committee Chair: Dr. Rory DePaolis

Committee Members/ Readers:

Dr. Yingjiu Nie

Dr. Geralyn R. Timler

Table of Contents

List of Tables

List of Figures

Abstract

Pupillometry holds great promise as a tool for infant language research but has not yet been used to probe word recognition. The goal of the described study was to design a functioning method that can later be used to test the possibility of using pupil dilation as a gauge of word recognition in 11-month-olds. To do this, we used the methods of an existing study (The Role of Accentual Pattern in Early Lexical Representation, Vihman, Nakai, DePaolis, & Hallé 2004) with modifications. Our method ran a one-sided head turn preference paradigm with the addition of an eye tracker for pupil data collection. Seven infant participants were tested, with adjustments made to the testing setup and data analysis methods after each. The literature provided little guidance for data analysis, so several analysis methods were attempted and found unsuccessful. Ultimately, it was found that maximum pupil dilation is generally greater following words a baby is likely to know than rare words. In addition, a baseline phase is necessary to establish meaningful criteria for comparison. We recommend that a full investigative study be performed using the methods and changes we have outlined here. This study has laid some of the groundwork for the investigation of pupillometry as a tool for infant language research.

Introduction

Pupillometry as a research tool has been in use since the 1970's, but until recently few researchers have applied it to infant studies (Hepach and Westermann, 2016). Even fewer have used pupillometry to probe infant language. The goal of the study was to create a workable method that can later be used to investigate pupil dilation as a reliable index of word recognition in 11-month-old infants. This was accomplished by using the methods of an existing study (*The Role of Accentual Pattern in Early Lexical Representation,* Vihman, Nakai, DePaolis, & Hallé, 2004) and adding a pupil dilation component. The referenced study assesses word recognition in 11-month-olds using the head turn preference paradigm (HPP). The current study used the same stimuli and similar procedures with the addition of an eye tracker to record the infant's pupil diameters throughout the process. The study designed a workable method that can later be used to test pupillometry's suitability or unsuitability as a tool for assessing word recognition.

Pupillometry Basics

The pupil is the opening in the eye through which light enters (Beatty & Lucero-Wagoner, 2000). Two sets of muscles control the size of the pupils, which primarily changes in response to light (Beatty & Lucero-Wagoner, 2000). An example of this is when someone's pupils constrict when walking into the bright sunshine from a dark building. In addition to light, pupils dilate during a cognitively difficult task and in response to new or stimulating information (Hepach & Westermann, 2016). Pupil dilation also occurs due to sensory stimuli (Qiyuan, Richer, Wagoner, & Beatty, 1985) like words or pictures. Pupil changes driven by cognition are smaller than light-driven changes, on the order of 0.5 mm (Beatty & Lucero-Wagoner, 2000). These changes are difficult to see with the naked eye, demonstrating the need for an eye tracker. The baseline diameter of the pupil is a response to the luminance of the environment, but changes compared to baseline give a “momentary, involuntary, and unbiased measure of arousal, attention, and cognitive load” (Sirois & Jackson, 2012). Because of this, researchers have proposed that pupillometry be used in infant studies as a companion or alternative to more traditional measures like looking time.

The increasing availability of eye trackers has made measuring pupil dilation more accessible to researchers. Tobii eye trackers are among the most common used in infant pupillometry studies (Sirois & Jackson, 2012). The eye tracker that was used for the current study was a Tobii TX300 model. In order for data to be taken, the infant must sit in front of the eye tracker screen and keep their eyes oriented towards the screen during calibration. Once calibrated, the eye tracker recorded data on where and how long the infant looks, in addition to the diameter of each pupil. This continued as long as the infant looked at the screen or until the experiment was concluded. Since pupil dimeter alone is sampled between 50 and 300 times per second (Hepach & Westermann, 2016), a wealth of data is obtained for researchers to use to infer cognitive effort. The eye tracker used in this study sampled at a rate of 60 times per second.

Why Pupillometry?

There are several factors that made pupillometry a promising tool for infant language research, and more specifically for this study. The first is that it avoids weaknesses of other methods. One of the most prevalent tools for studying infant speech perception, the head-turn preference paradigm (HPP), has been used with great success but has weaknesses that pupillometry does not. HPP and other tools used to measure infant speech perception depend on the infant performing a task, like turning their head or sucking more quickly on a pacifier (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995). If the baby is not in a mood to cooperate, or just is not paying attention, results from this test can be noisy. Looking times are documented to decrease when an infant becomes bored or tired, regardless of the stimuli (Hepach and Westermann, 2016, Vihman, Nakai, DePaolis, & Halle, 2004). To contrast, humans have little control over how and when their pupils dilate (Laeng, Sirois, & Gredebäck, 2012). All that the infant must do when pupil data is taken is look at the eye tracker screen. Pupillometry's ability to avoid weaknesses such as inattentiveness that affect tools like HPP suggest it has potential.

The second factor that made pupillometry an appropriate method for this study is that it's been successfully used in previous infant studies. Hepach and Westermann (2013) assessed 10 and 14-month-old infants' pupil diameters when viewing clips of actors interacting with a stuffed tiger. It was found that babies of both age groups displayed greater pupil dilation when viewing an actor with an angry expression gently petting the toy, an action incongruent with their expression, than when the same actor thumped the toy as would be expected from someone who is angry. Also, 14-month-olds, but not 10-month-olds, showed increased pupil dilation when viewing an actor with a happy expression thumping the tiger than petting it (Hepach & Westermann, 2013). This study suggested not only that infants are sensitive to the congruence of peoples' actions and emotions, but that reliable pupil data could be obtained from infants and analyzed to draw conclusions. Other researchers have also investigated infant cognition and emotion using pupil measurements (Jackson & Sirois 2009; Jessen, Altvater-Mackensen, & Grossmann, 2016; Jackson & Sirois 2012; Geangu, Hauf, Bhardwaj, & Bentz, 2011). One study used pupillometry alongside looking time measures as we did in this project (Jackson & Sirois, 2009).

Pupillometry has been used with infants not only in emotional perception, but also in language studies. Pupillometry has been used to investigate if 3 and 6-month-olds were sensitive to frequent vs. infrequent speech sounds (Hochmann & Papeo, 2014). They discovered that both age groups showed increased pupil dilation for infrequent vs. frequent sounds, suggesting that they were able to notice the difference. These researchers used the same method to show that 6-month-olds, but not 3-month-olds, recognized the same consonant in different syllables (Hochmann & Papeo, 2014). Thus, they stated that these older infants were able to solve the invariance problem, which was to recognize that these consonants were the same despite minor acoustic differences. Another study showed that infants from 9-14.7 months learned words more effectively when a show gesture was combined with the word. The evidence for this was both increased looking times towards the correct object in this condition, and increased pupil dilation (de Villiers Rader & Zukow-Goldring, 2015). Together, these studies showed that pupillometry could be successfully used to probe infant language. However, no study had yet been done that attempted to use pupillometry as an index of word recognition with babies.

Novelty and Familiarity

One of the major driving forces of infant language experiments is novelty and familiarity. Infants will notice either familiar or novel stimuli by increased attention, often shown by looking longer towards their preferred stimuli. The main factors controlling an infant's preference for new or familiar stimuli are the complexity of the stimulus and how long the exposure is (Mather, 2013). These factors interact in ways that aren't completely understood. In general, infants initially prefer familiar stimuli, but eventually switch to preferring novel ones (Rose et. al, 1982). This varies depending on the age of the infant, the complexity of the stimuli, and how the infant is habituated to it, but generally holds true (Hunter & Ames, 1983).

In addition, infants' pupils had been shown to dilate in response to novel stimuli (Hepach & Westermann, 2016). The more novel and significant the stimuli, the greater the dilation (Hepach & Westermann, 2016). This has been used to measure infant's responses to possible and novel events in a violation of expectations paradigm, combined with looking time data (Jackson & Sirois, 2009). The described experiment took advantage of this response by comparing the pupils' responses to what is known about infant's preferences for novel stimuli. The two often corroborated each other, which is discussed in more depth later in the paper. This finding demonstrated that the experimental setup was accurately measuring the infant's responses.

There are also parallel methods of measuring familiarity; such as a parent questionnaire. An example of a study that measured familiarity with a questionnaire is a pupillometry study on toddlers' responses to mispronounced words (Tamási, McKean, Gafos, Fritzsche, & Höhle, 2017). Thus, part of the experiment was to have the infants' parents indicate how familiar they believed their babies were with the words from the familiar list presented during the test, as reflected by their ratings on a Likert scale. The questionnaire was created by the researchers and is discussed in greater depth in the procedures section, as well as being included as Appendix 1.

Statement of Problem

Pupillometry has not been used to assess infant word recognition. Since pupillometry shows promise as a tool for infant cognition research, it should be possible to use it to investigate word recognition in this population. Since data from the HPP is inherently noisy, supplementing it with pupillometry could significantly advance our understanding of the formation of infant lexicon.

Hypothesis

We hypothesized that, as a group, the infants would show greater average pupil dilation in response to the words from the familiar list than the words from the rare list. This study aimed to develop the methods and procedures needed for future researchers to test this hypothesis.

Procedures

The study used the methods for experiment one of "The Role of Accentual Pattern in Early Lexical Representation", Vihman et al., 2004, with the addition of pupillometry. The original study used HPP to determine when infants began to show a preference for words they were likely to be familiar with versus phonetically matched unfamiliar words. Two lists of words and phrases were presented. One contained words and phrases that babies are likely to recognize, like "apple" and "thank you". The other contained rare words and phrases that infants are not likely to be exposed to, like “a noose” and “bridle”. The two lists are phonetically and phonotactically balanced to ensure that the infants are responding only to familiar words. A two-sided HPP procedure was followed, using the lists as stimuli. A group of 9-month-olds and a group of 11-month-olds were tested, all of whom were learning British English. 11 out of 12 11-month-olds listened longer to the list of familiar words than the list of rare words. Only 4 out of the 12 9-month-olds showed this pattern (Vihman et al., 2004). This suggests that, as a group, 11-month-olds prefer familiar words over unfamiliar, although 9-month-olds do not. In addition, the same words were used with 10 and 11-month-old babies who were learning American English. As a group, the 11-month-olds preferred the familiar list, while the 10-month-olds didn’t have a preference (R. DePaolis, personal communication, April 23, 2018).

Participants

Since the original study (Vihman et al., 2004) found a word recognition effect at 11 months, we recruited infants close to this age. A total of seven infants were tested, with a mean age of 12.07 months. The age range was from 10.2 to 13.6 months, with a median of 11.8 months. The infants were recruited from the Harrisonburg area through a mass email at JMU and posters in the community.

Stimuli

The stimuli were the same as in the original study except for the word “nappy” being replaced with “cookie”. This is because American infants are unlikely to know the word “nappy”, as the term for the object in the US is “diaper”. Even with this change, the lists remain phonetically balanced. The words and phonetic transcriptions are as follows (from Vihman et al., 2004):

Table 1: Words from Vihman et al., 2004

Table 1
Stimuli used in Experiments 1 and 2

	Familiar word	Phonetic transcription	Rare word (unfamiliar to infants)	Phonetic transcription
Trochaic	apple	/æpəl/	bridle	/braɪdəl/
	baby	/beɪbi/	cycle	/saɪkəl/
	button	/bʌtən/	fog light	/fɔglaɪt/
	mummy	/mʌmi/	maiden	/meɪdən/
	nappy	/næpi/	manna	/mænə/
	sleepy	/slipi/	mortar	/mɔtə/
	thank you	/θæŋkju/	thorough	/θʌrə/
Iambic	a ball	/əbɔl/	a bine	/əbaɪn/
	away	/əwėɪ/	a noose	/ənus/
	balloon	/bəlun/	compare	/kəmpɛə/
	fall down	/fɔldaʊn/	disturb	/dɪstɜb/
	tonight	/tənaɪt/	taboo	/təbu/

The sound levels were set using a sound pressure level meter to ensure the peak amplitude of the words was at 60 dB SPLA. A phonetic analysis reveals that the two lists have comparable characteristics of place, manner, and voicing of phonemes. The only noticeable difference is that the rare list has six palatal sounds, while the familiar list only has one. This reflects the greater number of /r/ phonemes in the rare list. The complete phonological analysis can be found in appendix 6.

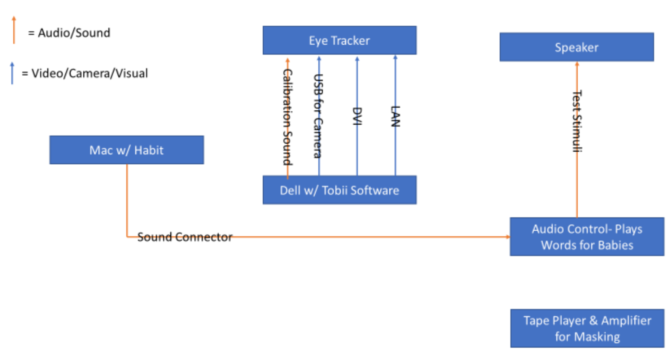
The words are pseudo-randomized into 12 lists of 12 words each, with each word appearing in the first or second position of one list. In addition, each 12-word list is split into three blocks of four words, with two familiar and two rare words in each block. This ensures that every infant hears every word at least once, and that they get an equal number of each type of list. Each list lasts between 24 and 26 seconds, with a 1.5 second gap between each word. The total length of all 12-word lists is 4 minutes and 56 seconds, plus 1-2 seconds for an attention getter between each list. This time between words was based on a review of the literature that suggested pupils would return to baseline from a cognitive-based dilation within this timeframe (Qiyuan et. al., 1985). At least one other study has used this interval successfully (Qiyuan et. al., 1985). While a longer interval might have better ensured the pupils return to baseline, it would risk losing the infant's attention. A shorter interval might not give enough time for a return to baseline. As an 11-month-old may not sit still long enough to complete all 12 trials, a participant's data will be considered useable if at least eight trials are completed. Even if only eight trials are finished, the order is still counterbalanced.

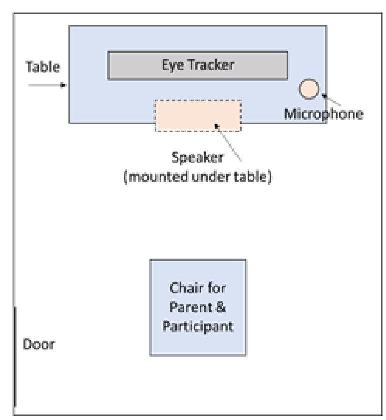
Procedure

This study used HPP as a supplement to pupillometry. This way, the results of the two measures could be compared to see if they corroborated each other. Even if the infant's pupils did not change, the researchers would still be able to assess word recognition using the HPP data. Other studies have successfully combined pupillometry with additional measures (Jackson & Sirois, 2009; Sirois & Jackson, 2012; Geangu1, Hauf, Bhardwaj, & Bentz, 2011). Thus, combining two measures like this had a successful precedent. The HPP procedure was slightly different from the original study. The original used a two-sided HPP, while this study used a one-sided HPP. This is because, in order for the eye tracker to collect data, the infant must be looking at it the majority of the time. A two-sided HPP requires the infant to look between two speakers, rather than straight ahead. A previous study used the word form recognition paradigm successfully with a one-sided HPP (Segal et. al., 2015). Thus, a one-sided HPP was a necessary change in order to accommodate the collection of pupil data.

To ensure the reliability and validity of the study, several steps were taken. The first is that the parent in the booth with the infant wore both insert earplugs and headphones playing masking babble. This kept the parent from hearing the words presented to their child and accidentally biasing the experiment. The observers also wore earplugs and remained blind to which words are playing. Observers were able to monitor the infant using the camera built into the eye tracker but could not hear any sound from the testing booth. The experiment was recorded using this camera and a small microphone in the booth. A naïve observer could then code the head turns in the videos to check reliability. A simplified block diagram of the setup is pictured below, as well as a diagram of the testing booth.

Figures 1 & 2: Block Diagram of Computer Setup (left) and Diagram of Testing Booth Setup (right)

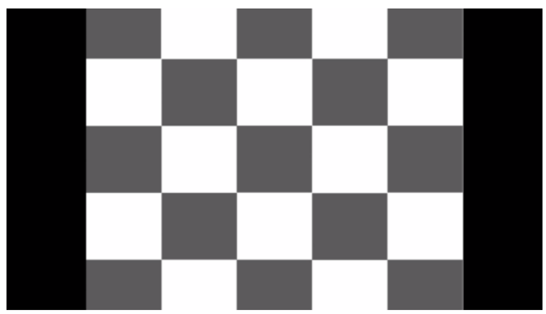




<u>Controlling Luminance</u>

The addition of pupillometry to the existing study means that care must be taken to control the luminance in the room. Even slight changes in environmental lighting could bias pupillometry data, as the recorded pupil dilation would not be due to the experimental stimuli but to the lighting in the room. This included not only overhead lights and lamps, but the brightness and contrast of computer screens (Hepach & Westermann, 2016). The eye tracker displayed a black and white checkerboard pattern throughout the test (see Figure 3). The checkerboard was static during word presentation but flashed by changing white squares to black and vice versa after a trial. Thus, the same image served as an attention getter and a neutral background. However, the luminance remained the same regardless of movement.

Figure 3: Visual Stimuli During Testing



As the project progressed, changes were made to better control the light levels in the booth. Initially, the only lighting in the booth was a dim overhead light. After the first participant, we became concerned that this was too much light and that a dimmer booth was needed to avoid biasing pupil data. For the second and third participants, the booth was lit by a dim floor lamp placed behind the participant's chair, as well as a small clip-on lamp overhead. This gave enough light to navigate the booth and code looks but reduced the overall brightness. The fourth, fifth and sixth participants were tested with a bright overhead light on, but no others. This change came from a discussion with a researcher who works with infant pupil measures, who suggested that a brighter room made pupil measures more accurate (G. Yao, personal communication, July 3, 2018).

This change also helped the researchers to code looks more easily, as it provided a better view of the infants' faces. In addition, the booth's white walls were covered in black fabric to reduce the contrast between them and the screen. This is discussed in more depth later in this paper.

Data Analysis

Pupillometry using an eye tracker is an emerging technique, and thus there is no gold standard for how to analyze data gained from experiments like the proposed study. There is great variation in variables such as when to start and stop measuring pupil dilation. In general, each author creates their own protocol for such analysis. A summary of relevant infant pupillometry studies and their analysis methods is presented in the following table.

Table 2: Overview of Data Analysis Methods from Literature

Study	**Author(s) & Publication Year**	**Used both L & R pupil or averaged**	**Analysis Method/Time Window Examined**	**Other Notes (Time between stimuli)**
Infants' sensitivity to the congruence of others' emotions and actions	Hepach, Westermann; 2013	Both pupil diameter values averaged and filtered for analysis	Pairwise comparisons, general linear mixed models fitted, time window not mentioned	No mention of pupil asymmetry, Time between video clips not mentioned, Looking time (not HPP) also measured
The Invariance	Hochmann, Papeo; 2014	Average change in pupils used	Experiment 1: significant effect in time window 883–2,183 ms,	No mention of pupil asymmetry,

Problem in Infancy: A Pupillometry Study		(not diameter), unclear if eyes were looked at individually or averaged	control was 0-883 ms, these parameters used in experiment 2, time window varied slightly between 3 and 6-month-olds	onsets of 2 consecutive syllables 750 ms apart in both experiments
Infant cognition: going full factorial with pupil dilation.	Jackson, Sirois; 2009	Filter was applied, then pupil diameters were averaged	ANOVA performed for looking times, Time window not mentioned	No mention of pupil asymmetry, Time between stimuli not stated
Pupillary responses reveal infants' discrimination of facial emotions independent of conscious perception	Jessen, Altvater-Mackensen, Grossmann; 2016	Pupil diameters averaged if both eyes available, considered for further analysis if not	Used Matlab, mean pupil diameter for entire trial calculated, controlled for individual differences by averaging separately for each participant and condition and dividing by overall mean pupil size of that participant	No mention of pupil asymmetry, Attention getter displayed for 1000 ms at end of trial
The Role of Speech-Gesture Synchrony in Clipping Words From the Speech Stream:	de Villiers Rader, Zukow-Goldring; 2015	Pupil diameter averaged	Analyzed time period where word occurred, period of same duration immediately before and after	No mention of pupil asymmetry

Evidence From Infant Pupil Responses				
Pupil Dilation and Object Permanence in Infants	Jackson, Sirois; 2012	Pupil diameters averaged, used one pupil to predict other's value if absent	Data filtered before analysis, window analyzed was entire trial	No mention of pupil asymmetry, time between stimuli not stated
Infant Pupil Diameter Changes in Response to Others' Positive and Negative Emotions	Geangu, Hauf, Bhardwaj, Bentz; 2011	Pupil diameters averaged, data interpolated from other eye if one missing	Data filtered before analysis, analyzed first 25s of stimulus presentation, 1 s baseline before stimulus onset used (baseline correction performed)	No mention of pupil asymmetry, 10 s attention getter between stimuli

All pupillometry data for this study was collected using the lab's Tobii TX300 eye tracker. The data was then exported to Microsoft Excel for analysis.

For the purposes of this project, the researchers initially used pupil average diameter for a 1/6 second (166.66 ms) period immediately before the onset of each word as a control. The main analysis window began one second/1000 ms after the onset of the word and lasted for 1/6 second (166.66 ms), again using average pupil diameter. These intervals were relatively small but provided plenty of data to work with as Tobii samples

60 times per second. 1/6 of a second is equal to ten data points on the Excel document that was used for data analysis. This interval was chosen based on a review of other infant pupillometry studies and knowledge of the pupillary system (Geangu, Hauf, Bhardwaj, & Bentz, 2011; Hochmann & Papeo, 2014; Qiyuan, Richer, Wagoner, & Beatty, 1985; Rader & Zukow-Goldring, 2015). It's known that pupils begin to dilate 0.3 to 0.5 seconds (300 to 500 ms) after stimulus, and that peak dilation occurs about 1 second later (Qiyuan et. al., 1985). Thus, it was believed that sampling the period where dilation peaks and comparing it to before the dilation starts would be effective. Ultimately, it was not, which will be discussed in greater depth later in the paper.

Not every word the infants heard was analyzed, but only the ones with the most complete data. For the familiar words, only the words rated most recognizable were analyzed, as it was believed that they would have the greatest pupil impact. Babies don't always learn words in the same sequence or at the same age, so the parent questionnaire showed us which words were truly familiar to each participant. Recognizability was measured with a parent questionnaire, which the researchers created. Parents rated their infants' recognition of each word on a Likert scale, with a rating of one indicating that the baby never recognized the word and a rating of 5 meaning that they always did. Most recognizable was defined as having a parent rating of 4 or 5 on a 5-point scale, with 5 indicating that the infant was believed to recognize the word every time. On rare occasions, a word with a 3 rating was used, but never a word with a lower rating. The words from the rare list were initially not included in the questionnaire, as they were intentionally chosen to be words that no infant would know. After the first infant participant, the words of the rare list were added to the questionnaire just in case the infant was exposed to them or similar-sounding words. A copy of the questionnaire is included as Appendix 1.

Interestingly, no studies have mentioned that pupils are not always perfectly symmetrical. While examining pilot data, as well as data taken from test runs of the experiment, it was noted that the right pupil was often slightly larger than the left. This difference wasn't large, only about 0.2 mm, but it was consistent enough to be noticeable. This also wasn't a cause for concern, as it's estimated that about 20% of the population has some degree of pupil asymmetry (Eggenberger, 2017). Considering that many authors treated the left and right pupils as interchangeable in terms of data collection, this presented a problem. To compensate for this, we averaged the diameters of the left and right pupils to create one value for analysis. This reduces the amount of data to analyze and prevents the differences between left and right pupil diameters from making the data uninterpretable. Six out of the seven studies summarized in the above table also averaged the left and right pupils, showing that averaging is a valid method for handling this kind of data.

Results from the First Participant

Initially, a single pilot participant was tested. This participant was 13 months old at the time of testing and was learning both Spanish and English. The pilot run provided a test of the experimental setup, demonstrating HPP and pupillometry can be measured simultaneously. The results of this analysis are as follows.

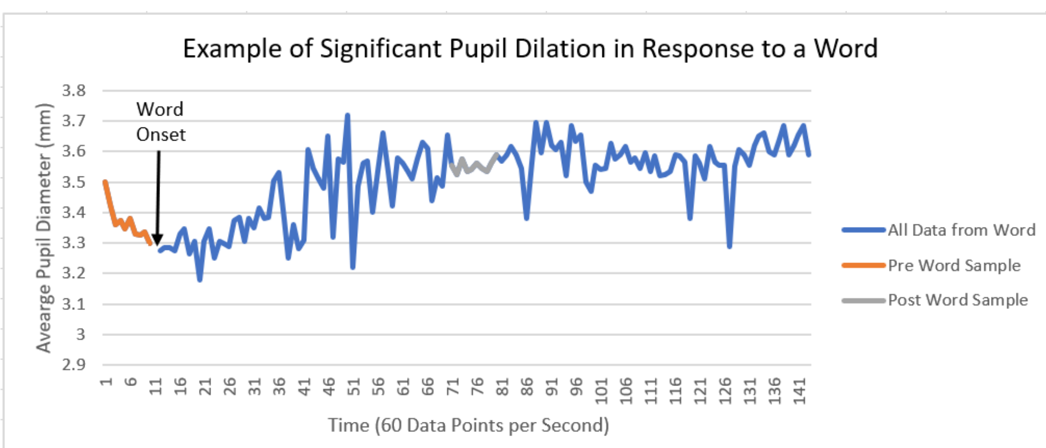
For the pupil dilation analysis, six familiar words and nine rare words from different trials were chosen for analysis because they had the most complete data, with either no or minimal missing data points. In addition, the analyzed familiar words were flagged by the mother on the questionnaire as very likely to be understood by the infant. The questionnaire form is shown in Appendix 1. The analysis was performed as described above, with the control and post word samples compared using T-tests. The results of this analysis are broken down by word type in the chart that follows.

Table 3: Pupil Data from Participant 1

Familiar Words			Rare Words		
Word	**Significant?**	**Constriction or Dilation?**	**Word**	**Significant?**	**Constriction or Dilation?**
Baby	Yes	Constriction	Compare	Yes	Dilation
Mummy	No	Dilation	Cycle	Yes	Constriction
A ball	Yes	Constriction	Manna	Yes	Dilation
Mummy	Yes	Dilation	A bine	Yes	Constriction
Baby	No	Constriction	Thorough	No	Constriction
Mummy	No	Constriction	Mortar	Yes	Constriction
			Manna	Yes	Dilation
			Mortar	No	Dilation
			Taboo	Yes	Dilation

An example of this analysis in graph form is as follows. More specifically, the graph represents word four from the familiar list, “mummy”, which had a statistically significant pupil dilation in response to the word.

Figure 4: Example of Significant Pupil Dilation in Response to a Word from Participant 1



It was hypothesized that infants would show greater pupil dilation in response to words from the familiar list than words from the rare list. The results of the first participant's data do not completely support this. There appears to be greater average pupil dilation in response to words from the rare list than the familiar list, at least in the words with enough data to analyze. It's possible that, since the participant is 13 months old rather than 11, the preference is for the novelty of the rare words over the familiarity of the recognized words. However, we later discovered that the pupil data from this subject was influenced by the testing booth itself, and thus our data from this participant are suspect.

The analysis of the HPP is more straightforward than that of the pupil data. Eleven trials were available for analysis, as the twelfth was not completed because of fussiness. This resulted in five rare and six familiar trials. Descriptive statistics were used to compare the mean looking times for each type of list. Both the mean and median looking times were greater for the rare lists than the familiar. The analysis is summarized in the table below.

Table 4: Participant 1 Head Turn Results

	Familiar	**Rare**
Mean (ms)	9296.8333	11182.4
StdDev (ms)	8122.737332	2530.21675
Median (ms)	6660	11472

Overall, the head turn data showed a slight preference for the rare words. The pupil dilation data appears to back this up, as there was greater average pupil change in response to the rare words than the familiar ones. This is different than what was hypothesized but may be due to the greater age and linguistic capabilities of the participant than was expected. Overall, the first participant showed that it was possible to run pupillometry measures side by side with the head turn preference paradigm. However, the pupil data was noisy and did not display a clear pattern.

Dilation and Constriction

While analyzing the data from the first participant, we were puzzled by the fact that the subject's pupils sometimes constricted after hearing a word rather than dilating. There was no apparent connection between constriction or dilation and the familiarity of a word. This was especially perplexing because there were no reports of constriction in the literature.

A careful examination of the testing booth alongside an inspection of pupil data and infant behavior revealed some testing issues. The walls of the testing booth were significantly brighter than the checkerboard patterned screen of the eye tracker. The baby looked between the screen and the walls repeatedly during the test as part of the head turn procedure. The difference in luminance between the darker screen and brighter walls caused the pupils to dilate and constrict depending on where the baby looked. The eye tracker registered these changes, which we then falsely attributed to the words the baby heard. Further review of the pupil data confirmed that each constriction was preceded by a look away from the screen. Because of this, the pupil data from the first participant was removed from the general analysis.

To keep the environment in the testing booth from influencing future pupil data, black fabric was hung over the white walls. This darkened the booth and prevented a subject's pupils from changing during a look away. A college aged person was used to test this, and this person's pupils showed no significant change while looking between the checkerboard and the darkened walls. Pictures showing what the booth looked like before and after are shown below. Subsequent participants were tested with the black fabric hung to foster reliable pupil data.

Figure 5: Testing Booth with White Walls

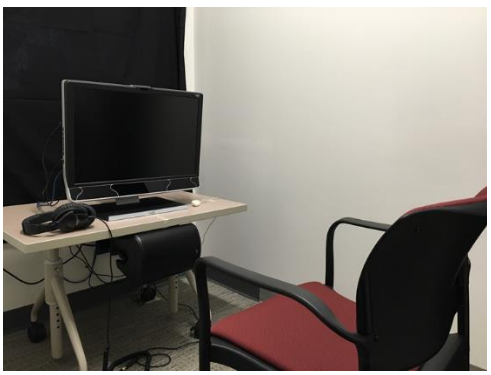


Figure 6: Testing Booth with Fabric Hung

<u>Results from the Second and Third Participants</u>

Participant 2 was a 13-month-old female. Due to fussiness, only six trials were completed. We analyzed this participant's data using the same methods as participant 1. With the white walls taken out of the equation, we believed that this would present a clear picture of how the baby's pupils responded to different words. However, the results were confusing. Every word, both familiar and rare, was followed by a statistically significant dilation (see table below).

Table 5: Participant 2 Significance by Word

Word #	Word	Parent Rating	P-value (2 Tail)	Pre Mean	Post Mean	Significant?	Constriction or Dilation?
1	Thank you 1	4	2.60126E-21	4.016	4.4575	Yes	Dilation
2	Sleepy	4	3.79419E-22	4.324	4.8525	Yes	Dilation
3	Cookie	4	4.14623E-22	4.176	4.8385	Yes	Dilation
4	Mummy	4	1.02835E-21	4.6245	4.836	Yes	Dilation
5	Thank you 2	4	2.02864E-19	4.24	4.7205	Yes	Dilation
6	Baby	5	6.55142E-22	3.214	3.761	Yes	Dilation
Word #	**Word**		**P-value (p<.05)**	**Pre Mean**	**Post Mean**	**Significant?**	**Constriction or Dilation?**
1	Foglight	1	4.14912E-22	4.032	4.598	Yes	Dilation
2	Maiden	1	6.72968E-21	3.512	4.226	Yes	Dilation
3	Taboo	1	4.20396E-09	3.8805	4.186	Yes	Dilation
4	Thorough	1	0.011078267	3.8595	4.041	Yes	Dilation

While the constriction issue was resolved, when analyzing the second participant's data according to protocols established from past studies, all of our tests for dilation produced statistically significant results. Participant 3, a girl 11 days short of 12 months old, exhibited similar results, with 21 out of 29 words analyzed showing a statistically significant dilation. Although not every word showed a significant dilation, the vast majority did. This pattern of results suggested our data analysis methods were suspect, despite their basis in the literature. For example, we were measuring a functionally meaningless change that did not indicate cognitive effort related to word recognition.

Since there were no answers in the literature, we experimented with different pre and post period lengths. Three familiar and three rare words were selected from participant 3's data, as they had the fewest missing values. The familiar words were chosen both for completeness and because they were rated most recognizable by the participant's parent. Five words were complete, with no missing values from the word's onset until the onset of the next word. The sixth word, "baby", had four missing values in that period. A description of the sampling periods used and a table of the results are presented below.

Table 6: Comparison of Sampling Periods for Selected Words from Participant 3

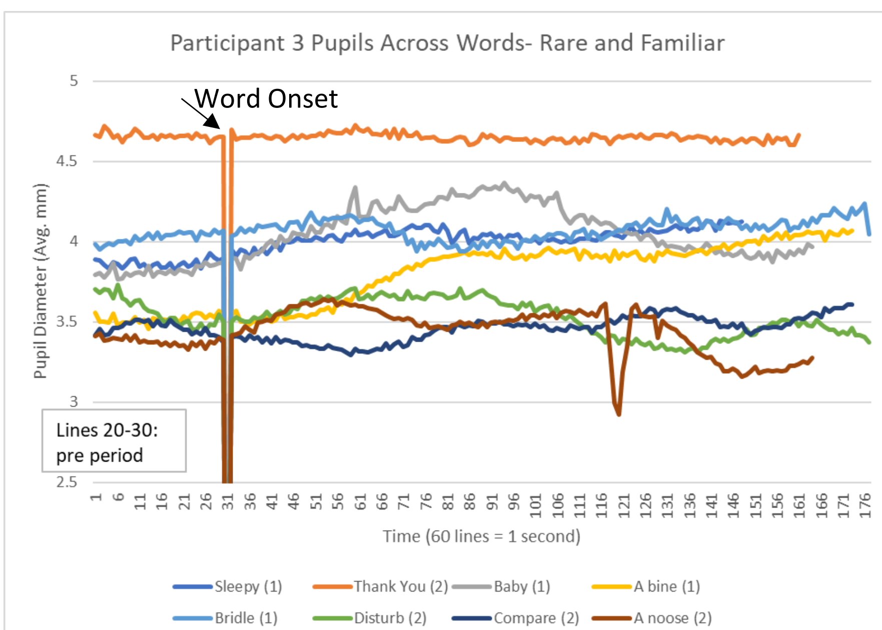
Word	Usual	0.5s 30LP	0.5 s 60 LP
Sleepy 1	Sig. Dilation	Sig. Dilation	Sig. Dilation
Thank You 2	No Sig. Change	No Sig. Change	Barely Sig. Constriction
Baby 1	Sig. Dilation	Sig. Dilation	Sig. Dilation
Bridle 1	Sig. Constriction	No Sig. Change	No Sig. Change
Compare 2	Sig. Dilation	Barely Sig. Constriction	No Sig. Change
A noose 2	Sig. Dilation	Sig. Dilation	Barely Sig. Dilation

Note: 60 lines = 1 second

Usual Sampling Period = 10 lines/0.16 s before word onset and 10 lines/0.6 s taken 60 lines after word onset
0.5 s sampling period 30 lines post = 30 lines/0.5 s before word onset and 30 lines/0.5 s taken 30 lines after word onset
0.5 s sampling period 60 lines post = 30 lines/0.5 s before word onset and 30 lines/0.5 s taken 60 lines after word onset

Although there were some instances where a longer sampling period meant fewer statistically significant changes, the data still did not show an interpretable pattern. We then plotted the words above, plus two others that were similarly complete, on a single graph (see below). This showed what happened with the baby's pupils across the course of each word, from 0.5 s before the word onset until the onset of the next word.

Figure 7: Pupil Diameter Across Words for Participant 3



The graph showed that the infant's pupils began and ended each word period at different diameters, suggesting the pupils were not returning to a baseline diameter between words. Also, there was no consistent pattern in how the pupils changed over the course of a word. Some periods presented a straight line with little change after a word, even words that the infant was believed to consistently recognize (e.g. "thank you"). Others showed a steady increase in pupil diameter (e.g. "a bine"). Most periods exhibited an increase and decrease in diameter with seemingly little relationship to the words presented (e.g. "compare"). There appeared to be a greater overall change over the course of familiar words than rare. Overall, examining whole word periods provided information we hadn't had before, but still didn't reveal an overall pattern until we started to consider the possibility that we were seeing a list effect for familiarity and novelty, rather than an effect for individual words.

Results from the Fourth and Fifth Participants

As the graph of pupil diameter across word periods for participant 3 demonstrated, an infant's pupils did not follow a consistent pattern after hearing a word. The changes in pupil diameter varied both in magnitude and timing. This made examining a brief pre and post word sample impractical, as there was no single time window where the change was greatest. We hypothesized that the maximum diameter reached would serve as a better metric, regardless of how long after the word onset it occurred. This was the approach we took to participant analyzing 4 and 5's data. Participant 4 was 10 months old at the time of testing, and fussy throughout. We did get some usable data, but less than with other subjects. Participant 5 was 11 months and three days old when tested. He was wiggly, but looked at the screen most of the time, allowing for good data collection.

At this point, we also reassessed our criteria for whether a word period was complete enough to analyze. Previously, any word period with long gaps where data was lost from both eyes was considered unusable. This criterion severely limited the number of analyzable word periods for each infant, especially if they were bored or fussy. However, not every data gap has the same cause. When the baby looks to the black walls of the booth, that produces a data gap that makes the word period unusable. The luminance of the walls is slightly different from that of the screen, so any pupil changes could be due to that rather than the words. If there's no change in what the baby looks at, that word period may still be analyzable. This can happen if the baby leans back into their parent but remains looking at the screen. The eye tracker records a video of the infant during the test, which can be reviewed to determine which type of data gap occurred. With this distinction in mind, several word periods that were previously thought too incomplete to analyze were reviewed on video. In cases where the baby's eyes remained on the eye tracker through the gap, the word was added to the analysis. This was especially helpful for participant 4, as there were no word periods without at least some data loss.

With this in mind, we determined the maximum diameter the infants' pupils reached in the period between one word's onset and the next. This analysis was done for all participants except the first, as participant 1's data was considered suspect.

Table 7: Maximum Pupil Diameter for Participants 2-5

	Mean		Median		Sig. Difference in Means?
	Familiar	Rare	Familiar	Rare	(p value)
P2	4.80125	4.32	4.8775	4.3	Yes (0.011599226)
P3	4.406667	3.854	4.365	3.71	Yes (0.039461804)
P4	5.111667	4.9875	5.025	4.9975	No (0.238316014)
P5	3.915	3.789444	3.95	3.79	No (0.192357664)

All four participants showed greater mean and median maximum pupil dilation following very recognizable familiar words than rare ones. The magnitude of the difference varied by participant. The difference was statistically significant for participants 2 and 3, but not for 4 or 5. Maximum dilation provided a pattern that held for all four participants with usable data, but the difference was difficult to quantify. The mean diameters could be compared between familiar and rare words, but there was still no value to use as a baseline diameter.

Results from the Sixth and Seventh Participants

It had become apparent that even though we followed guidelines from previous research, we were not getting a usable baseline for each infant's pupil diameters. This made it difficult to judge if a change had truly occurred. Thus, the addition of a baseline phase to the experimental protocol was necessary.

To establish a baseline measure for each infant's pupils, we added an interlude where the baby looked at the same static checkerboard that was used in the rest of the experiment, but with no words playing. This was placed at the beginning of the experiment, immediately after calibration, so that the infant had not yet been exposed to any of the test words. Since the same visual stimuli was used, there were no changes in luminance that could influence the infant's pupil size. To keep the baby's attention on the screen, an instrumental rendition of “Twinkle Twinkle Little Star” was played over the image. The baseline phase lasted until the infant oriented to the screen for at least two seconds, as judged by the researcher.

Two infants were tested in this final version of the protocol. The first, participant 6, turned 11 months old on the day of testing. The second, participant 7, was six days short of 12 months at the time of testing. Maximum pupil diameters were analyzed similarly to previous participants, with the results displayed below.

Table 8: Maximum Pupil Diameter for Participants 6 & 7

	Mean (mm)		Median (mm)		Sig. Difference in Means?
	Familiar	Rare	Familiar	Rare	(p value)
P6	4.78	4.49625	4.8575	4.455	Yes (0.032608603)
P7	5.035	5.57	5.145	5.57	No (0.125816406)

Participant 6 followed the pattern established with participants 2-5 in that the mean and median maximum pupil diameter was greater for familiar words than rare words. Like participants 2 and 3, the difference in means was statistically significant.

Participant 7 showed the reverse in that both the mean and median maximum diameter were greater for rare words than familiar. However, this must be interpreted with some caution. Data collection was compromised for this participant, as the eye tracker had difficulty finding the pupils. Even when the baby was looking directly at the screen, the pupils were not always captured. This resulted in very few usable words, with only five familiar and two rare words included in the analysis. The results from participant 7 were included for the sake of comparison but may be considered suspect.

In addition to maximum dilation, the data from the baseline phase was examined. This phase was used to establish how the participants' pupils responded when no words were being presented. Because the baseline phase length varied by participant, a small section was chosen for analysis. This section began two seconds (2000 milliseconds) after the beginning of the baseline period and lasted for half a second (500 milliseconds). There were no guidelines in the literature about the timing of a baseline sample, so this interval was chosen based on previous work on the study. The analysis period is highlighted in the figures below.

Figure 8: Pupil Diameters Across Baseline Period Participant 6

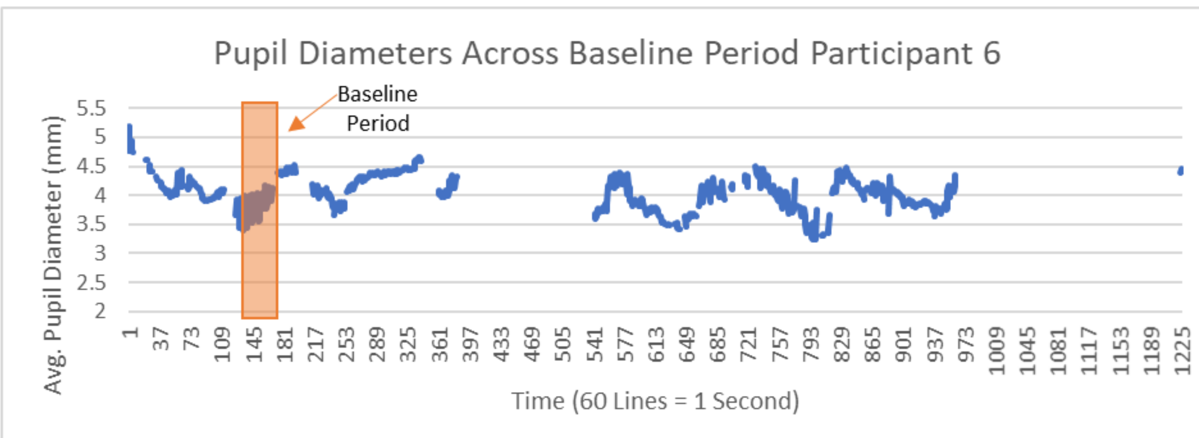
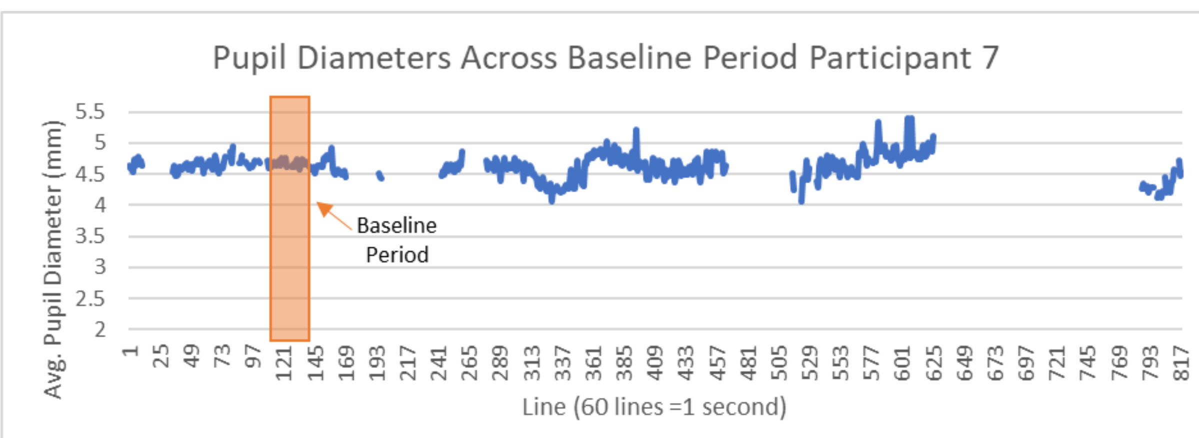


Figure 9: Pupil Diameters Across Baseline Period Participant 7



Participant 6 had a longer baseline period, but both displayed a similar pattern. The infant's pupils did not remain perfectly static across the baseline phase, even without luminance changes or other stimuli to affect them. Both participants' pupils displayed a wave-like pattern on the graph. This variation was more pronounced in participant 6, but also visible in participant 7. It is worth noting that although participant 7's data may still be considered suspect, it is much more complete during the baseline phase than while auditory stimuli were being played. This is likely due to the participant's greater focus on the screen initially, which decreased over time.

Head Turn Results

While pupil data was the main focus of the project, head turn data was also collected and examined. As described in the procedures, a one-sided head turn preference paradigm was run along with pupillometry measures. Data was collected by the Habit program and analyzed in Microsoft Excel. The results from all seven participants are summarized in Table 9 below. The mean looking times reflect each participant's times for rare and familiar lists, not for individual words within a list.

Table 9: Head Turn Results from All Participants

	Mean Looking Times (ms)		Familiar Greater?
	Familiar	Rare	
Participant 1	9296.8333	11182.4	No
Participant 2	21166.667	22419	No
Participant 3	15228.4	14696	Yes
Participant 4	5984	15756.6667	No
Participant 5	16424.833	16150.1667	Yes
Participant 6	9302.5	15210.6667	No
Participant 7	17161.667	13587.6667	Yes

It was expected that, as a group, the infants would have longer mean looking times for familiar lists than rare lists. However, the participants were different ages at the time of testing, which likely resulted in different language levels. Because of this, the group is too heterogenous for meaningful group comparisons to be made. Individual results were analyzed a test of the data collection setup but should not be considered representative of broader groups. On an individual level, three out of the seven participants followed the expected pattern of greater mean looking times for familiar lists than rare. Four out of the seven showed longer mean looking times for the rare lists. This may be because these infants were so familiar with the familiar words that they had begun to prefer the novelty of the rare words.

The main finding from the head turn portion of this project is that it's possible to run pupillometry and head turn measures simultaneously. The combination of the two measures provided more data about how participants responded to the words presented than either would have alone. This gave a more complete picture of the infants' word form recognition while building on previous infant language research. Overall, this study provided evidence that concurrently running head turn and pupillometry measures is both feasible and beneficial.

Conclusion

Pupillometry is an emerging tool that holds great promise for infant language investigation. This study outlined the development of a methodology that can later be used to assess pupillometry's potential as an index of word recognition in 11-month-olds. We began with an untested method, although there was evidence from other infant pupillometry studies that shows that the procedures and data analysis methods were reasonable. Each infant we tested revealed elements of the study that could be improved and forced us to fix problems that we didn't know existed. The final methodology is the result of all these adjustments and changes and has been used successfully to collect both pupil and head turn data.

The next step would be for another researcher to conduct a full experimental study to test the hypothesis we've presented. If another researcher chooses to take on a similar study, we have a few recommendations based on our experiences with the project.

- Use a more interesting visual stimuli than a plain checkerboard, possibly a moving figure to keep babies' attention while maintaining constant luminance
- Experiment with the baseline period- length and which section to analyze
- Create inclusion and exclusion criteria for when pupil data is usable
- Collect CDI data from participants to give more information on language development at time of test

If, as our data suggests, pupil dilation can be used to gauge word recognition, it would provide another tool to assess what babies know long before they can tell us verbally. We hope that the work we've done over the course of this project will encourage future researchers to undertake similar projects knowing that some of the challenges have already been resolved.

References

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (Second ed., pp. 142-162). Cambridge, United Kingdom: Cambridge University Press.

Eggenberger, E. R. “Anisocoria.” *MedScape*, 1 Dec. 2017, emedicine.medscape.com/article/1158571-overview.

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PloS One, 6*(11), doi:10.1371/journal.pone.0027132 [doi]

Hallé, P. A., & Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development, 17*(2), 119-129, doi:https://doi.org/10.1016/0163-6383(94)90047-7

Hepach, R., & Westermann, G. (2013). Infants’ sensitivity to the congruence of others’ emotions and actions. *Journal of Experimental Child Psychology, 115*(1), 16-29. doi:https://doi.org/10.1016/j.jecp.2012.12.013

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development, 17*(3), 359-377. doi:10.1080/15248372.2015.1135801

Hochmann, J. R., & Papeo, L. (2014). The invariance problem in infancy: A pupillometry study. *Psychological Science, 25*(11), 2038-2046. doi:10.1177/0956797614547918 [doi]

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology, 19*(3), 338-352. doi:10.1037/0012-1649.19.3.338

Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science, 12*(4), 670-679. doi:10.1111/j.1467-7687.2008.00805.x [doi]

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development, 18*(1), 111-116. doi:https://doi.org/10.1016/0163-6383(95)90012-8

Mather, E. (2013). Novelty, attention, and challenges for developmental psychology. *Frontiers in Psychology, 4*, 491. doi:10.3389/fpsyg.2013.00491 [doi]

Qiyuan, J., Richer, F., Wagoner, B. L., & Beatty, J. (1985). The pupil and stimulus probability. *Psychophysiology, 22*(5), 530-534. doi:10.1111/j.1469-8986.1985.tb01645.x

Rader, N. d. V., & Zukow-Goldring, P. (2015). The role of speech-gesture synchrony in clipping words from the speech stream: Evidence from infant pupil responses. *Ecological Psychology, 27*(4), 290-299. doi:10.1080/10407413.2015.1086226

Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology, 18*(5), 704-713. doi:10.1037/0012-1649.18.5.704

Segal, O., Keren-Portnoy, T., & Vihman, M. (2015). Infant recognition of Hebrew vocalic word patterns. *Infancy, 20*(2), 208-236. doi:10.1111/infa.12072

Sirois, S., & Jackson, I. R. (2012). Pupil dilation and object permanence in infants. *Infancy, 17*(1), 61-78. doi:10.1111/j.1532-7078.2011.00096.x

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language, 60*(2), 252-269. doi:https://doi.org/10.1016/j.jml.2008.11.003

Tamási, K., McKean, C., Gafos, A., Fritzsche, T., & Höhle, B. (2017). Pupillometry registers toddlers' sensitivity to degrees of mispronunciation. *Journal of Experimental Child Psychology, 153*, 140-148. doi:https://doi.org/10.1016/j.jecp.2016.07.014

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language, 50*(3), 336-353. doi:https://doi.org/10.1016/j.jml.2003.11.004

Appendix 1: Parent Questionnaire

Parent Questionnaire: Head Turn & Eye Tracking

Family Profile

Child's Name____

Birth Date__________________ Birth Place_______________________

	Mother	**Father**
Name		
Birthplace, date		
Accent		
How long have you lived in the valley?		
Occupation		
Highest level of education		

OTHER LANGUAGES to which child is exposed (indicate language, speaker and how often child is with speaker) ____

____

OTHERS living at home besides parents (indicate accent if relevant)

____

____

OTHER CARETAKERS (approx. amount of time per week spent with them; indicate accent) ____

____

____

DOES THE CHILD HAVE A HISTORY OF EAR PROBLEMS/INFECTIONS? Yes / No

If yes, explain: ____
____

Circle how well you think your baby recognizes these words/phrases. They may or may not have attached meaning to them yet, this just asks if they recognize the word if they hear it. **1 indicates they don't recognize the word at all, and 5 indicates that they always recognize the word.**

	1 (Never Recognizes)	**2**	**3**	**4**	**5** (Always Recognizes)
Away	1	2	3	4	5
Apple	1	2	3	4	5
Baby	1	2	3	4	5
Button	1	2	3	4	5
Cookie	1	2	3	4	5
Mommy	1	2	3	4	5
Sleepy	1	2	3	4	5
Thank you	1	2	3	4	5
Diaper	1	2	3	4	5
A Ball	1	2	3	4	5
Balloon	1	2	3	4	5
Fall Down	1	2	3	4	5
Tonight	1	2	3	4	5

Appendix 2: IRB Materials

James Madison University
Human Research Review Request

FOR IRB USE ONLY:				
Exempt:	☐	**Protocol Number:**	**1st Review:** ____	**Reviewer:** ________________
Expedited:	☐	**IRB:** 17-0541	**2nd Review:** ____	**Reviewer:** ________________
Full Board:	☒	**Received:** 04/07/17	**3rd Review:** ____	

Project Title:	Assessing Infant Word Recognition Through Head Turn Preference and Pupil Dilation
Project Dates: (Not to exceed 1 year minus 1 day)	From: 4/27/17 (MM/DD/YY) To: 4/26/18 (MM/DD/YY)
Responsible Researcher(s):	Amy Vinyard, Kierra Lynch
E-mail Address:	vinyaram@dukes.jmu.edu
Telephone:	(573) 823-2256
Department:	Communication Sciences and Disorders
Address (MSC):	4303
Please Select:	☐ Faculty ☐ Administrator/Staff Member ☒ Undergraduate Student (Kierra Lynch) ☒ Graduate Student (Amy Vinyard)
(if Applicable):	
Research Advisor:	Rory DePaolis
E-mail Address:	depaolra@jmu.edu
Telephone:	540-568-3869
Department:	Communication Sciences and Disorders
Address (MSC):	MSC 4303
Minimum # of Participants:	10
Maximum # of Participants:	20
Funding:	**External Funding:** **Yes:** ☐ **No:** ☒ **If yes, Sponsor:** _____ **Internal Funding:** **Yes:** ☒ **No:** ☐ **If yes, Sponsor: CSD Department** **Independently:** **Yes:** ☐ **No:** ☒
Incentives:	**Will monetary incentives be offered? Yes:** ☒ **No:** ☐ **If yes: How much per recipient?** $20 **In what form? Gift card**
Must follow JMU Financial Policy:	http://www.jmu.edu/financemanual/procedures/4205.shtml#.394IRBApprovedResearchSubjects

Institutional Biosafety Committee Review/Approval:	**Use of recombinant DNA and synthetic nucleic acid molecule research:** ☐ Yes ☒ No **If "Yes," approval received:** ☐ Yes ☐ No ☐ Pending **IBC Protocol Number(s):** **Biosafety Level(s):**
Will research be conducted outside of the United States?	☐ Yes ☒ No **If "Yes," please complete and submit the International Research Form along with this review application: http://www.jmu.edu/researchintegrity/irb/forms/irbinternationalresearch.docx.**
Certain vulnerable populations are afforded additional protections under the federal regulations. Do human participants who are involved in the proposed study include any of the following special populations?	☒ Minors ☐ Pregnant women *(Do not check unless you are specifically recruiting)* ☐ Prisoners ☐ Fetuses ☐ My research does not involve any of these populations
Some populations may be vulnerable to coercion or undue influence. Does your research involve any of the following populations?	☐ Elderly ☐ Diminished capacity/Impaired decision-making ability ☐ Economically disadvantaged ☐ Other protected or potentially vulnerable population *(e.g. homeless, HIV-positive participants, terminally or seriously ill, etc.)* ☒ My research does not involve any of these populations

Investigator: Please respond to the questions below. The IRB will utilize your responses to evaluate your protocol submission.

1. ☒ **YES** ☐ **NO** Does the James Madison University Institutional Review Board define the project as ***research?***

The James Madison University IRB defines "research" as a "*systematic* investigation designed to develop or contribute to *generalizable knowledge*." All research involving human participants conducted by James Madison University faculty and staff and students is subject to IRB review.

2. ☒ **YES** ☐ **NO** Are the human participants in your study ***living*** individuals?

"Individuals whose physiologic or behavioral characteristics and responses are the object of study in a research project. Under the federal regulations, human subjects are defined as: living individual(s) about whom an investigator conducting research obtains:

(1) data through intervention or interaction with the individual; or (2) identifiable private information."

3. ☒ **YES** ☐ **NO** Will you obtain data through ***intervention*** or ***interaction*** with these individuals?

"Intervention" includes both physical procedures by which data are gathered (*e.g.*, measurement of heart rate or venipuncture) and manipulations of the participant or the participant's environment that are performed for research purposes. "Interaction" includes communication or interpersonal contact between the investigator and participant (*e.g.*, surveying or interviewing).

4. ☒ **YES** ☐ **NO** Will you obtain ***identifiable private information*** about these individuals?

"Private information" includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, or information provided for specific purposes which the individual can reasonably expect will not be made public (*e.g.*, a medical record or student record). "Identifiable" means that the identity of the participant may be ascertained by the investigator or associated with the information (*e.g.*, by name, code number, pattern of answers, etc.).

---

5. ☐ **YES** ☒ **NO** Does the study present ***more than minimal risk*** to the participants?

"Minimal risk" means that the risks of harm or discomfort anticipated in the proposed research are not greater, considering probability and magnitude, than those ordinarily encountered in daily life or during performance of routine physical or psychological examinations or tests. Note that the concept of risk goes beyond physical risk and includes psychological, emotional, or behavioral risk as well as risks to employability, economic well being, social standing, and risks of civil and criminal liability.

CERTIFICATIONS:

For James Madison University to obtain a Federal Wide Assurance (FWA) with the Office of Human Research Protection (OHRP), U.S. Department of Health & Human Services, **all** research staff working with human participants must sign this form and receive training in ethical guidelines and regulations. "Research staff" is defined as persons who have direct and substantive involvement in proposing, performing, reviewing, or reporting research and includes students fulfilling these roles as well as their faculty advisors. The Office of Research Integrity maintains a roster of all researchers who have completed training within the past three years.

Test module **at ORI website http://www.jmu.edu/researchintegrity/irb/irbtraining.shtml**

Name of Researcher(s) and Research Advisor	Training Completion Date
Rory DePaolis	**2-23-15** Dr. DePaolis will complete the IRB Social/Behavioral Research Course – Refresher Course before the training expires
Amy Vinyard	**4-1-17**
Kierra Lynch	**1-28-17**

For additional training interests, or to access a Spanish version, visit the National Institutes of Health Protecting Human Research Participants (PHRP) Course at: http://phrp.nihtraining.com/users/login.php.

By signing below, the Responsible Researcher(s), and the Faculty Advisor (if applicable), certifies that he/she is familiar with the ethical guidelines and regulations regarding the protection of human research participants from research risks. In addition, he/she agrees to abide by all sponsor and university policies and procedures in conducting the research. He/she further certifies that he/she has completed training regarding human participant research ethics within the last three years.

____ ________________

Principal Investigator Signature Date

____ ________________

Principal Investigator Signature Date

____ ________________

Principal Investigator Signature Date

____ ________________

Faculty Advisor Signature Date

Submit an electronic version (in a Word document) of your **ENTIRE** protocol to researchintegrity@jmu.edu.
Provide a **SIGNED** hard copy of the Research Review Request Form to:
Office of Research Integrity, MSC 5738, 601 University Boulevard, Blue Ridge Hall, Third Floor, Room # 342

Purpose and Objectives

Please provide a lay summary of the study. Include the purpose, research questions, and hypotheses to be evaluated. (Limit to one page)

The purpose of this study is to assess whether pupil dilation is a reliable index of word recognition in infants. Previous studies have shown that 11-month-old infants recognize words independently of context (Halle et al 1994 and Vihman et al. 2004 and Swingley 2005). Traditionally, the head turn preference test has been used to assess word recognition in this population. However, this method relies on the infant's behavior and is influenced by lack of attention (due to teething for example). Pupil dilation is a physiological response that is independent of infant behavior. It's well-known that an adult's pupils dilate when they're working on a cognitively difficult task, and the same holds true for infants. However, no one has tested how word recognition influences an infant's pupil dilation. We hypothesize that infant's pupils will dilate in response to unfamiliar words, but not in response to familiar words. We will use the head turn preference paradigm as a parallel measure to assess concurrent validity.

Procedures/Research Design/Methodology/Timeframe

Describe your participants. From where and how will potential participants be identified (e.g. class list, JMU bulk email request, etc.)?

Participants will be monolingual English-learning 11-month-old infants from the Harrisonburg/valley area.

How will subjects be recruited once they are identified (e.g., mail, phone, classroom presentation)? Include copies of recruitment letters, flyers, or advertisements.

Recruitment will be through a bulk email to all JMU faculty, staff, and students and through flyers placed in the community (see attached).

Describe the design and methodology, including all statistics, IN DETAIL. What exactly will be done to the subjects? If applicable, please describe what will happen if a subject declines to be audio or video-taped.

The study will measure pupil dilation during the head-turn preference paradigm. Each method is described below.

Head Turn Preference Paradigm:

This will be a one screen head turn preference test. The infant sits on a parent's lap in a soundproof booth, facing out. The parent wears headphones that play a masking noise to prevent them from influencing the baby's responses. Directly across from the infant at their eye level is the Tobii screen. The screen displays a checkerboard pattern, which doesn't change throughout the experiment. The sound stimuli are presented at a comfortable level through a loudspeaker mounted under the computer. The infant is first habituated to the task, learning that a sound continues to be played as long as they look at the screen. The sound stops once they look away for two seconds or more, thus ending the trial. If the infant looks away for less than 2 seconds, the trial continues, but the looking-away time is not included in the length of look. Once this is established, the experimental phase begins. The infant is presented with two lists of words, one at a time and randomly ordered. One list contains words that the infant is likely to be familiar with, like "mommy" and "baby". The other contains words that the infant has probably not heard often, like "maiden". The process is the same for each of the 24 words (12 familiar and 12 unfamiliar). An attention-grabber animation is used to center the infant's attention on the screen. Once the infant is centered, a word from one of the lists is played through a speaker. The word is repeated until the infant looks away from the screen for longer than two seconds. Throughout this process, the Tobii eye tracker is recording the size of the infant's pupils in millimeters. A researcher watches through a two-way mirror to code the infant's looks. The researcher codes the duration of the infant's gaze towards the screen by pressing keys on a computer keyboard. The measurements are recorded using the Habit program. This process is repeated with each word list. All infant responses will be videotaped and checked for reliability by another researcher after testing is over.

Eye Tracker Method (concurrent with HPP):

Before the Head Turn Preference Paradigm begins, the Tobii will calibrate itself to the infant's eyes. This is accomplished by having the infant look at a dot onscreen as it moves into various positions. This will take less than a minute. During all phases of the Head Turn Preference Paradigm, the Tobii will measure and record pupil size using an infrared beam of light (the same light that a TV remote control uses).

Emphasize possible risks and protection of subjects.
The main potential risk is the exposure of subject information (i.e. names and ages). This will be mitigated by keeping all identifying information in a locked drawer in the locked Infant and Toddler Language Laboratory. All electronic data will only contain subject codes without names. The lab is also behind a door that requires swipe access. Since the infant will be his/her mother at all times, any issues related to emotional distress (surprise at the changing computer screen for example) will be a non-issue.

The other potential risk is that the near infrared light used by the eye tracker could cause seizures in people with photosensitive epilepsy. About 3-5% of people with epilepsy have this type, and it may happen even in people without a history of seizures. A person with

Photosensitive Epilepsy would also be likely to have problems with TV screens, some arcade games, and flickering fluorescent bulbs. To ensure that subjects are protected, the researchers will exclude any parent or infant with a diagnosis of epilepsy or a history of difficulties with TV screens or flickering lights from the study. This is explained in the consent form. We are also excluding anyone who uses a medical device that can be affected by infrared light.

What are the potential benefits to participation and the research as a whole?

Parents of participants will be engaged in a language task that emphasizes the parent role in word learning. Researchers in the field will potentially gain evidence that pupil dilation is a reliable index for infants' word recognition while acquiring a first language. In addition, the parent of each participant will receive a $20 gift card as compensation for their time.

Where will research be conducted? (Be specific; if research is being conducted off of JMU's campus a site letter of permission will be needed)

The research will take place in the Infant and Toddler Laboratory in room 5018 on the fifth floor of the Health and Behavioral Sciences Building.

Will deception be used? If yes, provide the rationale for the deception. Also, please provide an explanation of how you plan to debrief the subjects regarding the deception at the end of the study.

Deception will not be used.

What is the time frame of the study? (List the dates you plan on collecting data. This cannot be more than a year, and you cannot start conducting research until you get IRB approval)

The study will begin as soon as IRB approval is obtained, and will continue throughout the subsequent year.

Data Analysis

How will data be analyzed?

Pupil dilation data from the Tobii eye tracker will be analyzed using Matlab. Looking times will be collected from the Habit software and analyzed using SPSS. With both data t-tests will be run to determine if the dependent variables (pupil dilation and looking time respectively) are different between familiar and unfamiliar words.

How will you capture or create data? Physical (ex: paper or tape recording)? Electronic (ex: computer, mobile device, digital recording)?

Both previously mentioned computer programs capture data electronically. Tobi records the infants' eye movements on a screen as well as pupil dilation. It also takes a video of the parent and infant's faces, like a video from a video camera. Habit records when and how long the infants look towards a stimulus, in this case a word. We will also have a parent fill out a questionnaire regarding their infant's language development (see attached). The parent will also fill out another questionnaire, the MacArthur-Bates Communicative Development Inventories (CDIs), which assesses infant language. A copy of the MacArthur-Bates CDI is included at the end of the document.

Do you anticipate transferring your data from a physical/analog format to a digital format? If so, how? (e.g. paper that is scanned, data inputted into the computer from paper, digital photos of physical/analog data, digitizing audio or video recording?)
All pupil dilation and looking time data will be created digitally. The experiment will be videotaped for reliability using the Tobii software. Data from the questionnaire will be inputted into the computer from the paper questionnaires. Data from the MacArthur-Bates Communicative Development Inventories (CDIs) will also be inputted into the computer from paper questionnaires right away. All data entered into the computer will be de-identified.

How and where will data be secured/stored? (e.g. a single computer or laptop; across multiple computers; or computing devices of JMU faculty, staff or students; across multiple computers both at JMU and outside of JMU?) If subjects are being audio and/or video-taped, file encryption is highly recommended. If signed consent forms will be obtained, please describe how these forms will be stored separately and securely from study data.
The data will be stored on two computers in the CSD research labs on the fifth floor of the HBS building. The labs are behind a door that requires swipe access, and the labs themselves require another key. The computers are password protected. All digital data will only include participant codes. The video of the infants will be stored on an encrypted hard drive that is located in the Infant and Toddler Language Laboratory.

Who will have access to data? (e.g. just me; me and other JMU researchers (faculty, staff, or students); or me and other non-JMU researchers?)
Only the two principal investigators and the faculty advisor will have access to the participant names. De-identified data may be used for future student projects.

If others will have access to data, how will data be securely shared?
All data will be viewed in the Infant and Toddler Language Laboratory. Video will be stored on an encrypted hard drive.

Will you keep data after the project ends? (i.e. yes, all data; yes, but only de-identified data; or no) If data is being destroyed, when will it be destroyed, and how? Who will destroy the data?
All de-identified data will be kept on the same computers after the project ends. The paper surveys will be kept in locked cabinets in the Infant and Toddler Language Laboratory. The lab is behind a door that requires swipe access, and the lab itself requires another key. Video recordings will be kept in the same lab on encrypted hard drives. Participant information will be destroyed three years after the end of the project. This includes videos, which will be deleted from the encrypted hard drives three years after the end of the project. Paper data will be destroyed three years after the end of the project.

Reporting Procedures

Who is the audience to be reached in the report of the study?

The audience will be other researchers interested in infant language acquisition, as well as clinicians who work with infants.

How will you present the results of the research? (If submitting as exempt, research cannot be published or publicly presented outside of the classroom. Also, the researcher cannot collect any identifiable information from the subjects to qualify as exempt.)

The results of the study will be written up as a master’s thesis and an honor’s thesis, as well as published in a peer-reviewed journal. De-identified data may be used for future classroom instruction.

How will feedback be provided to subjects?

Feedback will not be provided to subjects during or after the experiment. The results of the study will be posted on the Infant and Toddler Language Laboratory website

Experience of the Researcher (and advisor, *if student)*:

Please provide a paragraph describing the prior relevant experience of the researcher, advisor (if applicable), and/or consultants. If you are a student researcher, please state if this is your first study. Also, please confirm that your research advisor will be guiding you through this study.

This is the first study for Amy Vinyard and Kierra Lynch. Both are advised by Dr. DePaolis, who has been studying infant language development for 25 years. Rory DePaolis, PhD, has been conducting experiments with human participants for thirty years, including either running or supervising a half dozen studies that have collected observational data from over 100 families in Wales, England, and the US. He has also run and/or supervised at least a dozen studies that have collected experimental data using the head turn preference paradigm.

Appendix 3: Informed Consent Forms

Parent Informed Consent: Head Turn & Eye Tracking

Identification of Investigators & Purpose of Study

Your child is being asked to participate in a research study conducted by Amy Vinyard and Kierra Lynch from James Madison University, under the advisement of Dr. Rory DePaolis. This study is designed to establish if pupil dilation is a reliable measure of word recognition in 11-month-old babies. We will be running all parts of the study at James Madison University, Harrisonburg, Virginia, USA. The experiment will be videotaped.

Research Procedures

Should you agree to allow your child to participate in this research study, you will be asked to sign this consent form once all your questions have been answered to your satisfaction. You will also be asked to visit the James Madison University Infant and Toddler Language Laboratory. While your infant is seated on your lap, different types of speech will be presented through loudspeakers. Your child's response to this speech will be observed and videotaped, with your permission. An instrument called an eye tracker will use an invisible infrared light to measure the size of your baby's pupils. The presentation level of the speech will be about that of normal conversational speech. You will be asked to wear headphones playing noise and to use insert earplugs to mask the speech your infant is hearing so that your response does not affect your infant's response.

This study also consists of a questionnaire that will be administered to individual participants in the Speech Laboratory at James Madison University. You will be asked to provide answers to a series of questions related to the language of your child.

Time Required

Participation in this study will require 30 minutes of you and your infant's time.

Risks

The investigators do not perceive more than minimal risks from you or your infant's involvement in this study.

Benefits

Potential benefits from participation in this study include learning more about the way that infants begin to learn and remember their first words. Researchers in the field will potentially gain evidence that pupil dilation is a reliable index for infants' word recognition while acquiring a first language.

Incentives

You will be paid a $20 gift card for your participation.

Confidentiality

The results of this research will be presented at conferences and in the classroom. The results of this project will be coded in such a way that the respondent's identity will not be attached to the final form of this study. The researchers retain the right to use and publish non-identifiable data. All data will be stored in a secure location accessible only to the researchers. Upon completion of the study the data will be archived on non-networked digital media and stored in a secure laboratory.

There is one exception to confidentiality we need to make you aware of. In certain research studies, it is our ethical responsibility to report situations of child abuse, child neglect, or any life-threatening situation to appropriate authorities. However, we are not seeking this type of information in our study nor will you be asked questions about these issues.

Participation & Withdrawal

Your infant's participation is entirely voluntary. You may withdraw your infant from the study at any time. Should you choose to participate, you can withdraw at any time without consequences of any kind.

Questions about the Study

If you have questions or concerns during the time of your infant's participation in this study, or after its completion or you would like to receive a copy of the final aggregate results of this study, please contact:

Dr. Rory DePaolis

Communication Sciences and Disorders

James Madison University

depaolra@jmu.edu

(540) 568-3869

Amy Vinyard

Communication Sciences and Disorders Graduate Student

James Madison University

vinyaram@dukes.jmu.edu

Questions about Your Rights as a Research Subject

Dr. David Cockley

Chair, Institutional Review Board

James Madison University

(540) 568-2834

cocklede@jmu.edu

Giving of Consent

I have read this consent form and I understand what is being requested of my infant as a participant in this study. I freely consent for my infant to participate. I have been given satisfactory answers to my questions. The investigator provided me with a copy of this form. I certify that I am at least 18 years of age.

I give consent to be videotaped during my participation. (*yes/no*)__________ (parent's initial)

I give consent for use of my video in classrooms and conferences. (*yes/no*) _____ (parent's initial) This is not necessary to participate in the study.

____

Name of Child (Printed)

____

Name of Parent/Guardian (Printed)

____________________________________ _____________

Name of Parent/Guardian (Signed) Date

____________________________________ _____________

Name of Researcher (Signed) Date

Appendix 4: Recruitment Email Text

Email subject line:

Infant Scientists Wanted

Email Text

The Infant and Toddler Language Laboratory is in need of participants to study early language development. If you have an infant 12 months old or younger, consider participating in studies on how babies learn language.

In our laboratory at JMU, we are currently running a study to investigate pupil dilation in response to the presentation of words (IRB #17-0541). Your baby will sit in your lap and watch a video screen while the size of his/her pupils are measured in response to words presented over speakers. Each study on campus takes about 30 minutes, with your infant's portion lasting about 10 minutes. Your infant never leaves your arms. You will be compensated for your time.

If you are interested, please e-mail (vinyaram@dukes.jmu.edu) or call the Infant and Toddler Language Laboratory at (540) 568-8886. We appreciate your consideration in having your child(ren) participate in our studies!

Please feel free to share with others who may be interested.

Appendix 5: Recruitment Poster

HEY PARENTS!

Is your baby 12 months old or younger?

JMU's Infant and Toddler Language Lab needs your help!

If you have an infant 12 months of age or younger, please contact our laboratory to contribute to a study on how babies learn language! We are especially interested in babies who are 11 months old or within a few weeks of it.

IRB # 17-0541

Please visit our website:

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

JMU Infant & Toddler Language Lab
vinyaram@dukes.jmu.edu
(540) 568-8886

CONTRIBUTE TO A STUDY ON HOW BABIES LEARN LANGUAGE

YOUR BABY NEVER LEAVES YOUR LAP

STUDIES TAKE ABOUT 30 MINUTES

ONLY REQUIRE ONE VISIT TO THE JMU CAMPUS

YOU WILL BE COMPENSATED FOR YOUR TIME

FOR MORE INFORMATION, PLEASE CONTACT:

AMY VINYARD

vinyaram@dukes.jmu.edu

(540) 568-8886

JMU Infant & Toddler Language Laboratory

Appendix 6: Phonological Analysis of Words Presented

	Consonants														
	Place							*Manner*						*Voicing*	
Familiar	Bilabial	Alveolar	Velar	Labiodental	Interdental	Palatal	Glottal	Stop	Nasal	Fricative	Affricate	Glide	Liquid	Voiced	Voiceless
Apple	1	1						1					1	1	1
Baby	2							2						2	
Button	1	2						2	1					2	1
Mummy	2								2					2	
Cookie			2					2							2
Sleepy	1	2						1		1			1	1	2
Thank You		1	1		1	1		1	1	1		1		2	2
A Ball	1	1						1					1	2	
Away	1											1		1	
Balloon	1	2						1	1				1	3	
Fall Down		3		1				1	1	1			1	3	1
Tonight		3						2	1					1	2
Totals	10	15	3	1	1	1	0	14	7	3	0	2	5	20	11
	Place							*Manner*						*Voicing*	
Rare	Bilabial	Alveolar	Velar	Labiodental	Interdental	Palatal	Glottal	Stop	Nasal	Fricative	Affricate	Glide	Liquid	Voiced	Voiceless
Bridle	1	2				1		2					2	4	
Cycle		2	1					1		1			1	1	2
Fog Light		2	1	1				2		1			1	2	2
Maiden	1	2						1	2					3	
Manna	1	1							2					2	
Mortar	1	1				2		1	1				2	3	1
Thorough					1	1				1			1	1	1
A Bine	1	1						1	1					2	
A Noose		2							1	1				1	1
Compare	2		1			1		2	1				1	2	2
Disturb	1	3				1		3		1			1	3	2
Taboo	1	1						2						1	1
Totals	9	17	3	1	1	6	0	15	8	5	0	0	9	25	12